

0. Datos de interés

La totalidad del proyecto de título “Aplicación web y análisis de señales acústicas provenientes del llanto de los infantes (A web-based application and acoustic signal analysis of the infant cry)” fue realizada entre los meses de Febrero y Junio del año 2007 en el departamento de ingeniería de control e informática de la universidad “Budapest University of Technology and Economics” de Budapest.

El tutor de proyecto en la universidad de Budapest fue Mr. Várallyay György, siendo co-tutorado por el profesor de la universidad Carlos III de Madrid, Don Carlos Bousoño.

1. Introducción

El proyecto realizado en la universidad ‘Budapest University of Technology and Economics’ consiste en el desarrollo del software necesario para la segmentación y extracción de las principales características de las señales provenientes del llanto de los recién nacidos.

Adicionalmente fue implementada una aplicación web con el propósito de facilitar y almacenar archivos de audio con los correspondientes llantos para su posterior análisis y procesamiento. Esta información es convenientemente guardada posteriormente en una base de datos diseñada específicamente para esta aplicación.

Para la implementación del proyecto se utilizaron dos ordenadores Pentium IV con procesador a 3GHz. El software utilizado para el desarrollo de la aplicación fue el siguiente:

- Matlab R16 7.1: Utilizado para la implementación de los algoritmos de segmentación y extracción de las principales características de las señales.
- Eclipse SDK 3.3: Utilizado para el desarrollo del software de la aplicación web.
- Apache Tomcat 5.5: Servidor instalado en el PC y utilizado para las pruebas de la aplicación web.
- PSPad Editor: Utilizado para el desarrollo de la página web de la aplicación.
- MySQL Server 5.0: Base de datos utilizada para almacenar los archivos procedentes de la aplicación.

El proyecto consta de 5 capítulos principales. Una introducción que detalla el contenido del proyecto. Un segundo capítulo con el desarrollo del mismo. El tercero dedicado a las conclusiones. Un cuarto adjuntando el código fuente desarrollado, y finalmente, el quinto capítulo con el presupuesto.

2. Descripción del trabajo

La descripción del proyecto está dividida en dos apartados principales. La primera incluye el estudio y desarrollo de la aplicación web, mientras que la segunda hace lo propio con respecto al análisis de las señales.

2.1. Aplicación web

El estudio del arte de las aplicaciones web analiza en profundidad las posibles tecnologías que pueden ser utilizadas para el desarrollo de la aplicación. Así, se explican y valoran técnicas como CGI (Common Gateway Interface), las Java Servlets, las JavaServer Pages, y las utilizadas para el desarrollo y programación de las bases de datos, tales como JDBC.

Posteriormente se procede a la descripción de la aplicación web desarrollada. El objetivo principal de la aplicación web es la obtención del máximo número de muestras de llantos posible para su posterior análisis y extracción de características. De esta manera, será posible obtener mayor fiabilidad a la hora de clasificar y obtener conclusiones sobre posibles deficiencias auditivas de los bebés responsables de los llantos analizados.

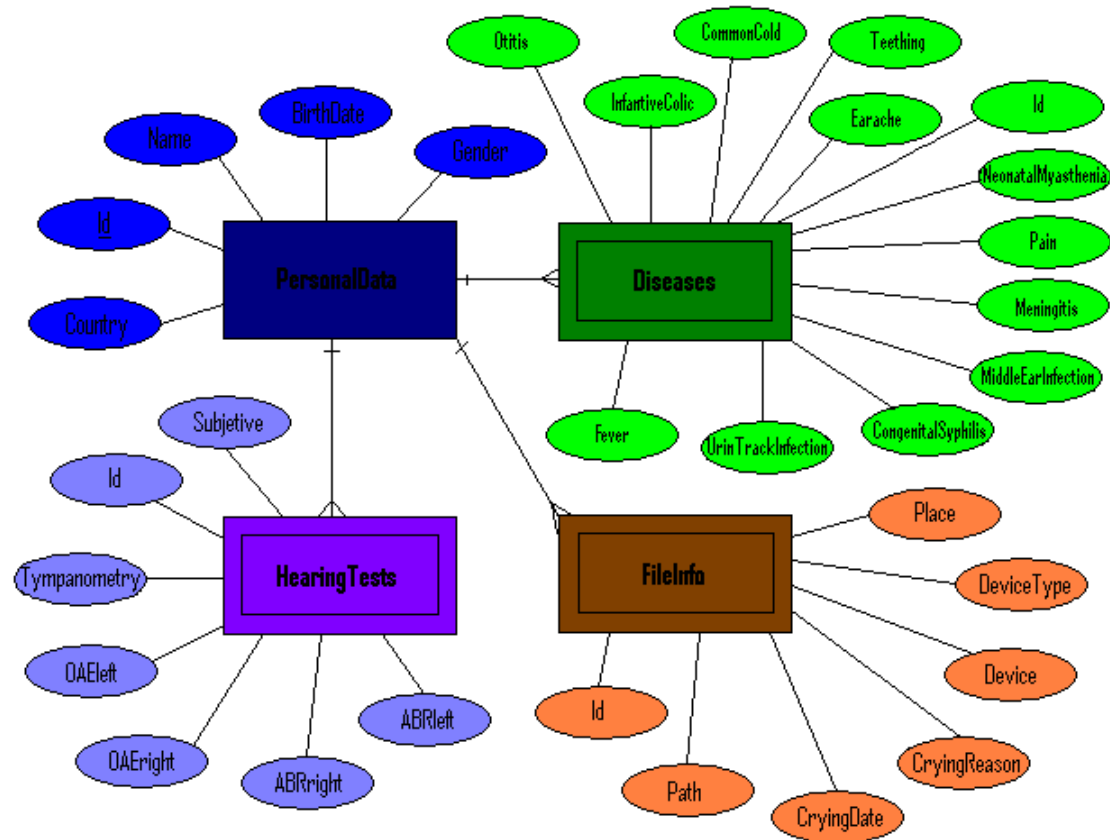
Para ello, se desarrollo la página web que se incluye en el anexo de este resumen. Dicha página contiene un formulario que incluye los aspectos más importantes sobre la identidad, enfermedades, tests auditivos y entorno en el que el llanto fue grabado. Esta información es relevante a la hora de extraer las singularidades del llanto y su posterior clasificación. Adicionalmente, se incluye un campo en el cual el fichero del llanto puede ser adjuntado y enviado.

Por último, la página web está diseñada de tal modo que el usuario tiene la posibilidad de escoger el idioma en que el formulario es mostrado en el lado del cliente, pudiendo elegir entre inglés, español o húngaro.

Cuando toda esta información es enviada al servidor web, éste se encarga de separar los datos concernientes al formulario del archivo de audio, así como almacenarlo convenientemente en la base de datos.

La base de datos como se ha comentado fue creada específicamente para esta aplicación. A continuación se muestra la estructura de la misma:

Crying files Data Base



El motor de la base de datos usado es InnoDB, cuya principal mejora con respecto a otros motores de almacenamiento disponibles para ser usado con MySQL es que soporta la llamada ACID-compliant transacción así como la integridad declarativa referencial.

2.2. Procesamiento de señal

Este apartado de la memoria del proyecto incluye el estudio del arte sobre las diferentes técnicas usadas para el procesamiento de señales de voz, tales como 'Automatic Speech recognition' (ASR), análisis de voz y segmentación de la misma, y métodos VDA para realizar esta segmentación.

Posteriormente se procede a la descripción de los principales parámetros de las señales de llanto que serán obtenidas. Estas son:

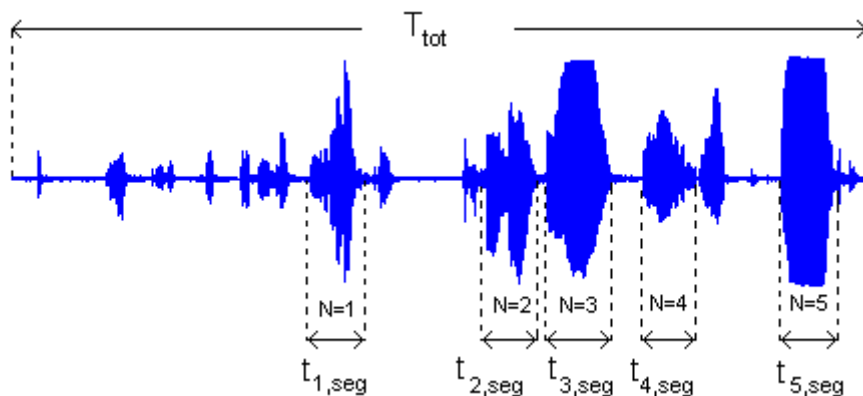
- Longitud media del segmento

- Longitud máxima del segmento
- Longitud del segmento específico de llanto.
- Frecuencia fundamental

Estos parámetros vienen íntegramente determinados por tres propiedades de la señal:

- Longitud de la señal (T_{tot})
- Longitud del segmento ($T_{seg} = \sum_N t_{i,seg}$)
- Número de segmentos (N)

Estas propiedades se muestran en la siguiente imagen:

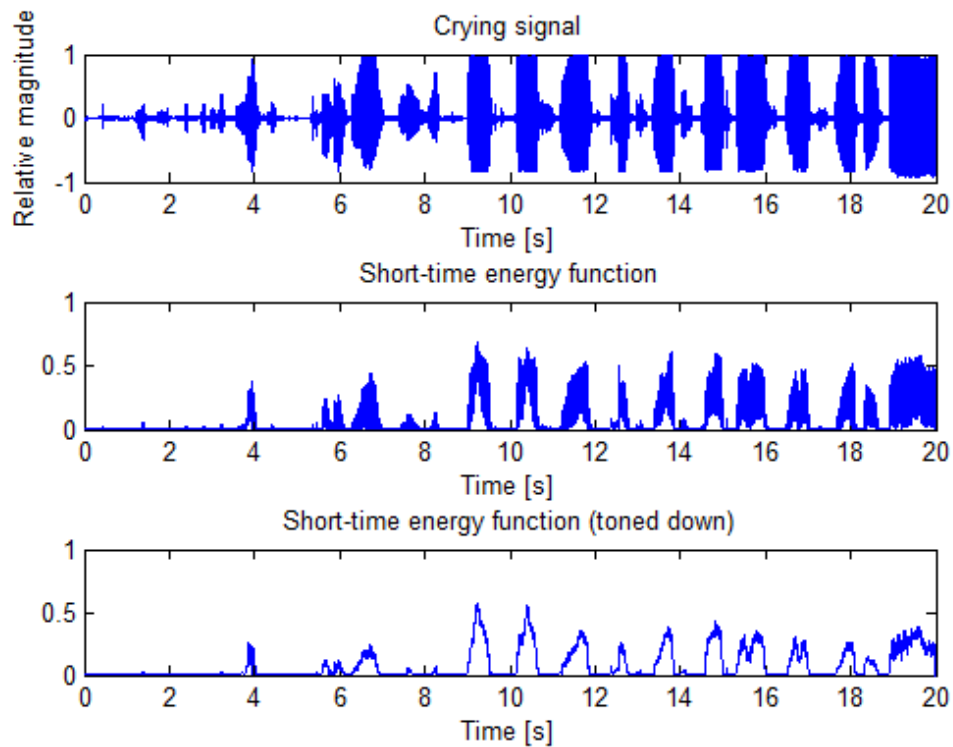


Para aislar la señal de llanto pura con efectos no deseados tales como ruido de fondo, tos, eco, y para su posterior segmentación, se implementaron tres algoritmos diferentes: segmentación por energía media, cruce por ceros y en el dominio de la frecuencia.

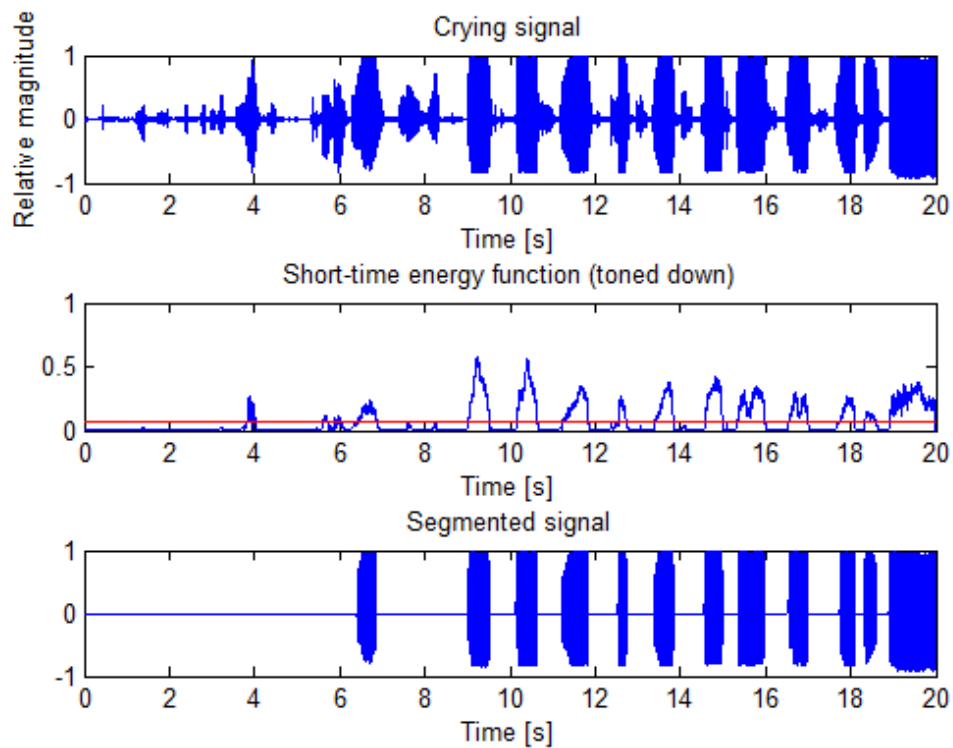
En la memoria del proyecto cada uno de estos algoritmos es explicado desarrollado y testado para señales de llanto de diferente calidad de grabación. Para los tres algoritmos, la señal resultante después de ser procesada según las particularidades del algoritmo, es pasada por un filtro de altas frecuencias con el fin de atenuar las variaciones rápidas y evitar la aparición de falsos segmentos fruto de pasos rápidos de la señal por el umbral calculado durante el procesamiento de la misma en una ventana de pequeña longitud.

A continuación se muestran los resultados obtenidos al procesar, atenuar y segmentar las señales dependiendo del algoritmo utilizado:

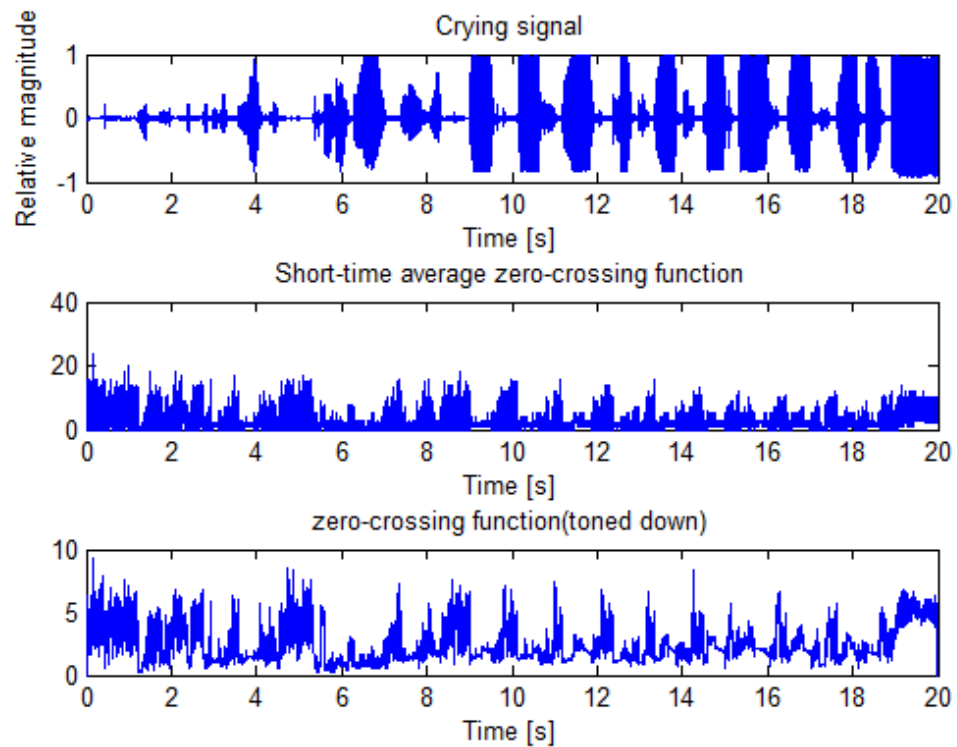
Utilizando el algoritmo de energía media:



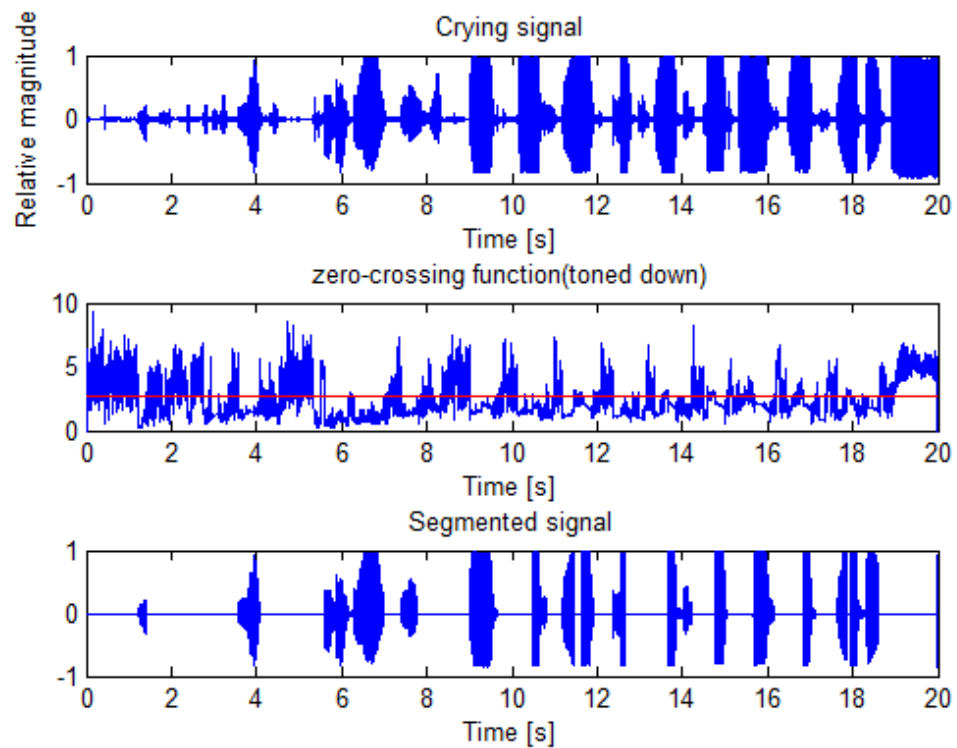
La siguiente figura muestra el resultado de la segmentación en función de un umbral calculado como un 10% por debajo del valor de la energía media de la señal:



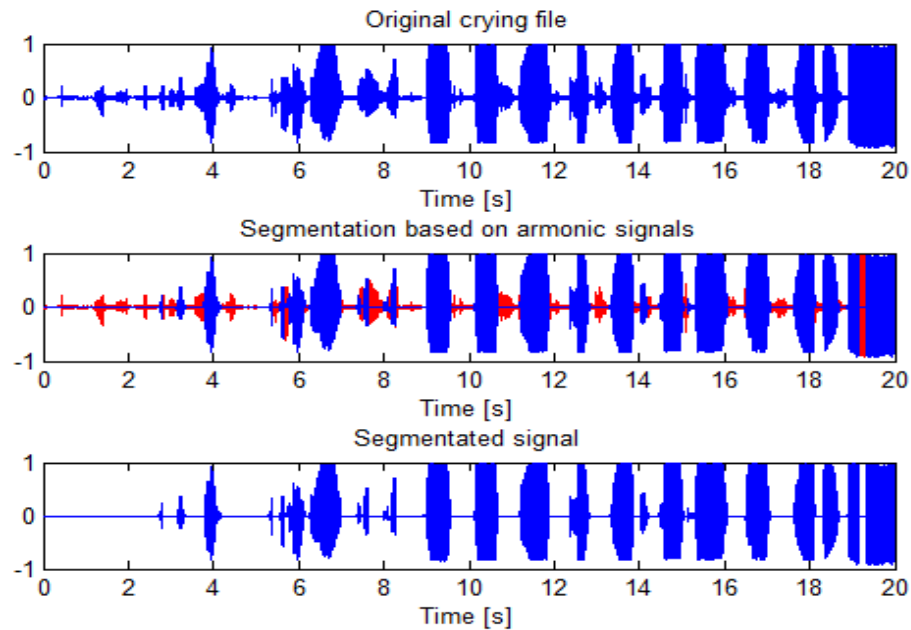
Para el algoritmo de cruce por cero:



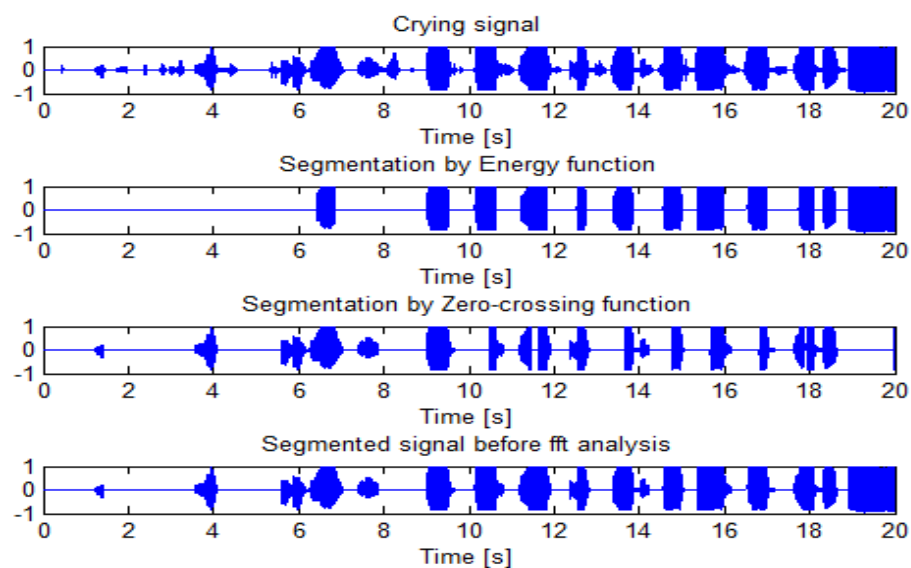
En este caso, el umbral tomado para la segmentación es calculado como un 10% por encima del valor de la media de cruces por cero de la señal:



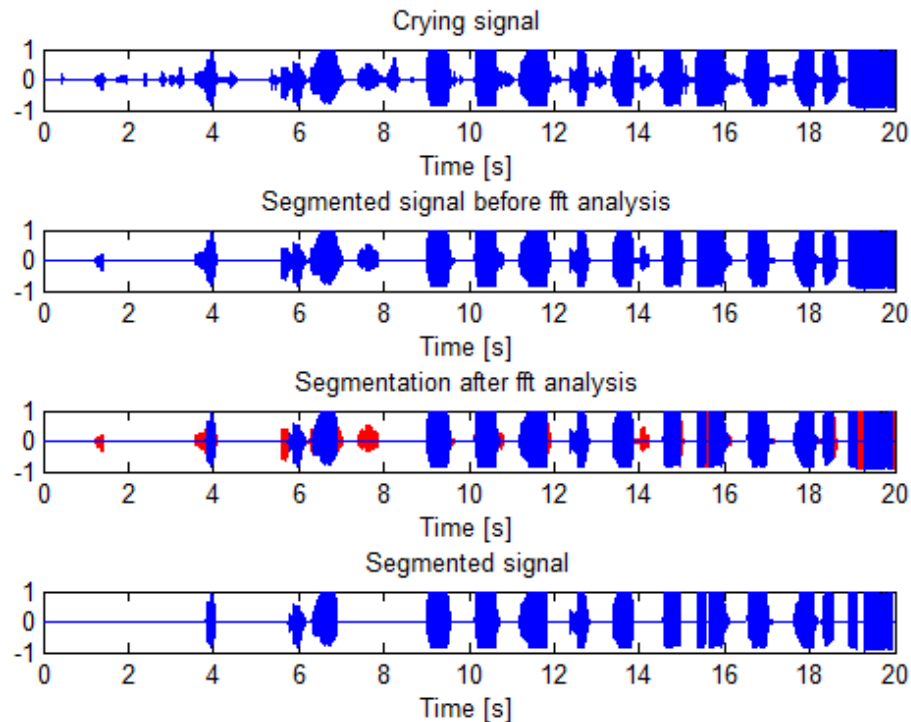
Finalmente, se utilizó un algoritmo de segmentación en el dominio de la frecuencia basado en el reconocimiento de patrones armónicos en la señal al ser procesada en intervalos de pequeñas ventanas. A continuación se muestra el resultado:



El algoritmo de segmentación que se decidió utilizar para la implementación del proyecto está basado en la utilización de las tres técnicas descritas anteriormente. Esta decisión fue tomada basándose en que cada algoritmo era óptimo para señales de diferente calidad. Así, para buscar generalidad en los resultados, la segmentación se realiza superponiendo los resultados de los dos primeros algoritmos y pasando la señal resultante por el filtro en el dominio de la frecuencia. A continuación se muestra el resultado antes de filtrar usando la FFT:



El resultado final de la segmentación para la futura clasificación de la señal es:



3. Conclusiones

A tenor de los resultados obtenidos tras testear el modelo implementado con los ficheros muestra suministrados por el departamento, se pudo demostrar que el modelo era válido y capaz de identificar diferencias comunes entre las muestras provenientes de bebés sanos y las que provenían de bebés con deficiencias auditivas.

Sin embargo, los resultados no son totalmente concluyentes debido a la insuficiencia de número de muestras suministradas. Por este motivo, se desarrolló la idea de crear una aplicación web para poder reunir el mayor número de ficheros posible y así poder ajustar con mayor precisión el modelo en un futuro.

4. Código fuente

En este capítulo se adjunta todo el código generado tanto para la aplicación web como para el procesamiento de la señal de audio.

5. Presupuesto

En este capítulo se detallan los costes de material y personal que fueron necesarios para el desarrollo del proyecto. La cantidad total asciende hasta un valor de 25.955 euros.

ANEXO (Página web desarrollada para la aplicación web)

English
Hungarian
Español

Fields marked with an asterisk* are required

Personal Info

Name: John Doe

*Gender: male ☒ female ☐

Country: Australia

*Date of birth : 05 08 2006

Birth weight (g): 3200

Diseases

☐ Common Cold

☐ Congenital Syphilis

☐ Earache

☒ Fever

☐ Infantive Colic

☐ Meningitis

☐ Otitis

☐ Urinary Track Infection

☐ Teething

☐ Middle Ear Infection

☐ Neonatal Myasthenia

☒ Pain

Hearing Tests

Auditory Brainstem Response (ABR): 0-20 right ear

0-20 left ear

Otoacoustic Emissions (OAEs): Response right ear

Response left ear

tympanometry : right ear

left ear

subjective test : Reaction

Crying file

Place of the recording hospital room

* Type of the recording device digital camera

Recording device

eg:Sony DSC-T100, Nokia N93...

Reason of crying (if known) [select]

* Date of crying 18 11 2007

* Attach crying file here C:\Documents and Settings\Examinar...

Submit